

Historical perspective on the development of outcomes measures for low-vision and blind rehabilitation in the Department of Veterans Affairs

Judith L. Babcock-Parziale, PhD;^{1-2*} Michael D. Williams, PhD³⁻⁴

¹*Southern Arizona Department of Veterans Affairs (VA) Health Care System, Southwestern Blind Rehabilitation Center, Tucson, AZ;* ²*Department of Psychology, University of Arizona, Tucson, AZ;* ³*Rehabilitation Research and Development Center of Excellence for Aging Veterans with Vision Loss, Atlanta VA Medical Center, Decatur, GA;* ⁴*Center for Assistive Technology and Environmental Access, Georgia Institute of Technology, Atlanta, GA*

Abstract—This article reviews the instrument development process and synthesizes current research findings for three low-vision and blind rehabilitation outcomes measures developed in the Department of Veterans Affairs (VA). The examined measures include the Blind Rehabilitation Service Functional Outcomes Survey, the Functional Assessment of Self-Reliance on Tasks, and the VA Low Vision Visual Functioning Questionnaire. We examined vision rehabilitation literature using two fundamental measurement criteria to compare and contrast the instrument development process and research findings for the three measures. Our findings suggest the three measures need refinement to meet the two criteria and proposed measurement standards for instruments in this field. Advanced development goals for the instruments are identified. Measures that meet the specified and proposed measurement criteria will help establish an evidence-based system and guide practice at the VA and in the field.

Key words: aging, blind rehabilitation outcomes, clinician ratings, low-vision outcomes, measurement, outcomes assessment, psychometrics, Rasch model, self-report ratings, visual impairment.

INTRODUCTION

This article reviews the instrument development process and synthesizes current research findings for three low-vision and blind rehabilitation outcomes measures

developed in the Department of Veterans Affairs (VA). A recent article by Robert Massof discusses the measurement of visual disability over the past 20 years [1], but this article reports on the state of vision rehabilitation outcomes measurement as it pertains to efforts supported

Abbreviations: ADL = activities of daily living, BRC = blind rehabilitation center, BRFIM = Blind Rehabilitation Functional Independence Measure, BROS = Blind Rehabilitation Outpatient Specialist, BRS = Blind Rehabilitation Service, BRSFOutSur = Blind Rehabilitation Service Functional Outcomes Survey, CR = clinician-rated, CTT = classical test theory, FAST = Functional Assessment of Self-Reliance on Tasks, FIM = Functional Independence Measure, FIMBA = Functional Independence Measure for Blind Adults, HRQOL = health-related quality of life, IADL = instrumental activities of daily living, NEI VFQ = National Eye Institute Visual Functioning Questionnaire, RR&D = Rehabilitation Research and Development, SR = self-rated, SWBRC = Southwestern Blind Rehabilitation Center, VA = Department of Veterans Affairs, VA-13 = Blind Rehabilitation Service Functional Outcomes Survey, VA LV VFQ = VA Low Vision Visual Functioning Questionnaire.

*Address all correspondence to Judith L. Babcock-Parziale, PhD; Southern Arizona VA Health Care System, Southwestern Blind Rehabilitation Center (3-124), 3601 South 6th Avenue, Tucson, AZ 85723; 520-792-1450, ext 5698; fax: 520-629-4995. Email: judith.babcock@med.va.gov
DOI: 10.1682/JRRD.2005.07.0125

by the VA Rehabilitation Research and Development (RR&D) Service and the VA Blind Rehabilitation Service (BRS). This review provides a historical overview of the impetus for developing the outcomes measures, an examination of the instrument development process, and a synthesis of the research findings based on a contemporary psychometric analytic approach.

METHODS

We examined low-vision and blind rehabilitation literature, VA RR&D Service project reports, and related conference abstracts and presentations to assemble a historical synopsis of instrument development efforts. From the synopsis, we identified three VA low-vision and blind rehabilitation outcomes measures and examined them for differences and similarities in the following areas: purpose of the instrument, content prioritization, content selection and item generation, item responses, field testing, and administration time and approach. We then evaluated research findings for the three instruments. Our efforts to synthesize the findings were somewhat restricted by the lack of designated standards for evaluating an instrument in this field. We therefore chose to apply the following two fundamental measurement criteria that most psychometricians believe outcomes instruments should satisfy to ensure optimal measurement properties [2]:

1. A low-vision and blind rehabilitation outcomes measure must reliably discriminate between the different levels of functional ability imposed by visual impairment. This criterion mandates that outcomes measures capture functional ability across the broadest array of functional domains. Instruments that meet this criterion will be able to measure patients' visual functioning across the continuum of services provided by the VA BRS.
2. A low-vision and blind rehabilitation outcomes measure must be sensitive to or able to reliably detect change across time for all levels of visual functioning.

These criteria merely reiterate two measurement properties—an instrument must discriminate both between and within veterans and do so accurately over time [2].

RESULTS

Department of Veterans Affairs Blind Rehabilitation Program

For more than five decades, the leadership, programs, and principles established within the VA BRS have contributed significantly to raising the level of quality services for severely visually impaired veterans in the United States and abroad. Veterans are admitted to an inpatient blind rehabilitation center (BRC) when a treatment cannot reverse visual impairment or when a decrease in visual function results in impairment (e.g., degenerative eye diseases or irreversible trauma). Veterans admitted to a BRC are legally blind (i.e., visual acuity of 20/200 or worse) or have a visual field diameter of 20° or less as measured by Goldmann perimetry.* The VA blind rehabilitation program is designed to improve the quality of life for severely visually impaired veterans by helping them develop the skills and abilities needed for personal independence, emotional stability, and successful integration into the community and family environments. In essence, the primary treatment goal is to make it easier for visually impaired persons to perform everyday activities [3].

In addition to providing quality services, the VA has taken a lead role in outcomes research in the field of low-vision and blind rehabilitation. The need to account for low-vision and blind rehabilitation program outcomes is crucial since the veteran population averages some 10 to 15 years older than the nonveteran U.S. population [4]. This age difference, coupled with the age-related nature of severe visual impairment, may cause the veteran population to show peak rates of vision loss a decade earlier than the nonveteran population. As a result, outcomes measures and low-vision outpatient services developed in the VA could potentially serve as models for nonveteran agencies as they respond to the dramatic increase in patients with severe visual impairment that is likely to occur over the next decade.

*We use the term "low vision" to describe those legally blind individuals who demonstrate usable residual vision and can benefit from some form of vision-enhancement training. The term "blind" refers to patients who have little to no residual vision and require vision-substitution training, which focuses on the use of senses other than vision.

The Beginning of a New Era

Beginning in 1995, changes in the Veterans Health Administration required that all inpatient and outpatient services provide evidence of treatment outcomes [5]. Ophthalmology and optometry rely on clinically objective measures such as visual acuity to provide an assessment of a patient's visual status. These measures, however, may not reflect the degree of visual impairment the patient experiences in his or her daily activities nor do they provide the information necessary to evaluate the functional gains acquired through the vision rehabilitation process [1,6]. While a growing recognition exists of the importance of patient-reported outcomes of vision rehabilitation, a lack of suitable measures also exists. In fact, Raasch et al. reported in 1997 that low-vision research has been hampered by a lack of standardized measurement tools [7]. In response to this growing concern, VA vision rehabilitation researchers began to identify strategies that address the lack of suitable outcomes measures for the field. Strategies ranged from evaluating current measures to discussing the development of new ones.

Evaluation of Vision-Specific Instruments

Vision rehabilitation investigators were initially challenged to identify measures that would account for low-vision and blind rehabilitation treatment outcomes at the 10 VA inpatient BRCs. In 1996, the 51-item National Eye Institute Visual Functioning Questionnaire (NEI VFQ) was introduced as a vision-specific measure of health-related quality of life (HRQOL) [8–9]. Soon after the NEI VFQ was introduced, two separate VA studies were conducted to evaluate the newest version, the 25-item NEI VFQ (NEI VFQ-25), as a measure of relevant outcomes following low-vision and blind rehabilitation. Babcock-Parziale and Head administered the NEI VFQ-25 to more than 200 veterans at admission and 6-weeks postdischarge from the VA's Southwestern BRC (SWBRC) in 1998. They found that the directions were not suitable for use with severely visually impaired individuals who often use visual aids before and after treatment. In fact, since the directions neglected to include the phrase "using your vision aids," no treatment effect was detected.* Stelmack et al., at the Edward Hines Jr. VA Hospital BRC outside Chicago, conducted a pilot study

with 77 veterans using the interviewer-administered version of the NEI VFQ-25 [10–11]. The investigators determined that almost one-quarter of the items on the NEI VFQ-25 were not relevant to persons with severe visual limitations. As a result of these pilot studies and others conducted outside the VA [3], a growing consensus in the field was that the full benefits of low-vision and blind rehabilitation services would go undetected unless visual functioning instruments were carefully designed to measure appropriate and relevant patient outcomes.

The assertion that new instruments were indeed necessary to measure outcomes was confirmed directly by Margolis et al.'s 2002 literature review of vision-specific instruments [6]. The authors examined the development and psychometric properties of 22 vision-specific instruments used to assess visual functioning and/or the impact of visual impairment on HRQOL or daily activities. Most of the instruments were developed to assess treatments for eye conditions such as cataracts, which means that patients' visual acuity was restored close to normal or correctable by lenses or contacts. The review indicates that the 22 vision-specific instruments are not routinely used to assess low-vision and blind rehabilitation outcomes because the instruments are intended to measure restored vision, *rather than levels of vision loss or disability* (emphasis added).

An additional concern about existing vision-specific questionnaires has to do with the assumptions underlying the Likert scoring used in the rating scales. Massof was the first published investigator in the field to question the assumptions underlying Likert scoring of visual function questionnaires [1,12]. He maintained that questionnaires such as the NEI VFQ-25 that rely on raw scores to determine the amount of rehabilitation change are not measures because one cannot assume that respondents' ratings produce ordinal scales. To qualify as a measurement, the instrument must produce an interval or ratio scale [1,12–13].

Intensive Item Analysis

VA instrument developers recognized the need to create measures with true interval scales. As a result, investigators began to analyze scales and items with greater care, an approach that McKnight et al. refer to as "intensive item analysis."† Intensive item analysis is a

*Babcock JL, Head DN, Massof RW, Stelmack JA. Low vision rehabilitation outcome measures estimated from a self-assessment questionnaire depend on patient instructions. In: Proceedings of the International Conference on Measurement; 2001 Nov; Chicago, IL.

†McKnight PE, Babcock-Parziale JL, Head DN. Ten essential steps to developing a blind rehabilitation outcome measure. Unpublished observations, 2006.

process of discovery that yields a set number of patterns that are largely explained by item difficulty (i.e., a scale value given to a particular item). Using this approach, the instrument developer carefully culls the data looking for patterns that are likely to represent a cluster of items with similar difficulty levels or probability of endorsement. Anomalous findings can suggest random responding, poor items, strange patterns by respondents, or a whole host of other circumstances that threaten the construct validity of the results. Several widely accepted methods (e.g., latent response models and latent variable models) and tools (e.g., item response theory, Rasch modeling, and confirmatory factor analysis) allow for a more intensive analysis of items [1–2].*

An example of the intensive item analysis approach is demonstrated by Stelmack et al. in their use of the Rasch person-item map to analyze the field-test version of the VA Low Vision Visual Functioning Questionnaire (VA LV VFQ) [14]. This analytic approach provided the instrument developers information regarding the range, precision, construct validity, and reliability of the instrument. While a review of Rasch modeling is beyond the scope of this article, this analytic approach and its application to low-vision and blind rehabilitation measures are well documented [1,3,11–14].*

Development of Department of Veterans Affairs Sponsored Instruments to Assess Visual Functioning

Over the last decade, we have been involved in the development of three instruments to measure the visual function outcomes associated with low-vision and blind rehabilitation services. In response to increasing demands for a functional outcomes measure for the BRS, the Atlanta VA RR&D Center hosted a meeting of the BRS Task Force on Outcome Measures in May 1996. Members of the appointed BRS Task Force included four BRC Directors and four VA vision rehabilitation researchers. The goal of the BRS Task Force was to develop a self-report outcomes measure that could be used to measure the effectiveness of blind rehabilitation. With the assistance of De l'Aune and colleagues at the Atlanta VA RR&D Center, the BRS Functional Outcomes Survey (BRSFOutSur) was developed [15–16]. The BRSFOutSur initially had 48 items, but after revisions, the number

of items was reduced and the instrument was renamed the VA-13.

Researchers at the SWBRC participated in the initial meetings of the BRS Task Force. They recognized that no standardized clinical assessments existed in the field to determine the efficacy of a BRS inpatient program; therefore, in 1997, the investigators began developing a clinical measure. Hence, the second instrument developed was the Functional Assessment of Self-Reliance on Tasks (FAST), an 11-item clinician-rated (CR) scale developed by Drs. Head and Babcock-Parziale at the SWBRC in Tucson, Arizona [17–20].

The third instrument is the VA LV VFQ, originally intended to serve as a low-vision supplement to the NEI VFQ. This 48-item self-rated (SR) scale was developed by Dr. Stelmack at the Hines BRC [3,14,21–22]. A description of each instrument and findings from the initial field tests are available in the referenced citations as well as **Appendices 1–3** (available online only at www.rehab.research.va.gov). The purpose of this article, as previously noted, is to highlight the similarities and differences across the three measures, cull the findings, and advance a theory of measurement in the field. We began by examining the instrument development steps: purpose of the instrument, development approach taken, content prioritization, item selection process, generation of item responses (or scales), approach and timing for administration of instruments, analytic strategies, and finally, instrument revision process and future research goals. **Table 1** illustrates the similarities and differences in the approaches used to develop the instruments.

Purpose of the Instruments

The first step in the development of a measure should be the identification of the purpose of the instrument [20].* The purpose of the measure dictates the task of development—the number of items, the range of content, and the type of response scale. Specifying the intended use of the measure is also extremely important because it determines the type of data necessary to collect as well as from whom to collect the data; hence, every instrument developed requires some thought about whom to sample. In some cases, an instrument is used for decision-making purposes (e.g., to treat or not to treat or to determine who has passed or failed some type of training). Instruments used for decision-making purposes are often referred to as “screens” in the psychometric literature.

*McKnight PE, Babcock-Parziale JL, Head DN. Ten essential steps to developing a blind rehabilitation outcome measure. Unpublished observations, 2006.

Table 1.

Similarities and differences across instrument development process.

Instrument	Purpose	Development Process	Initial Content Selection	Item Responses	Administration Time (min)
FAST	Clinical efficacy; program outcomes	Measurement practice	Empirical (Delphi)	10-pt Likert/ Independence	15
VA-13	Effectiveness	Measurement practice	Theoretical	3-pt Likert/ Assistance	15–20
VA LV VFQ	Effectiveness	Measurement development	Empirical (Delphi)	4-pt Likert/ Difficulty	25–30

FAST = Functional Assessment of Self-Reliance on Tasks, VA-13 = Blind Rehabilitation Service Functional Outcomes Survey, VA LV VFQ = Department of Veterans Affairs Low Vision Visual Functioning Questionnaire.

A screen refers to the ability of an instrument to order items. If an instrument cannot order respondents, but does a reasonable job at discriminating between two meaningful levels of responses, it is considered a screen [23]. The term “screen” commonly refers to an instrument that is easily administered in a short period of time and provides more limited information than a measure does about the level of a person on some specific characteristic. Screens provide the greatest amount of information around the point of discrimination, whereas measures must provide information over the entire range of potential response levels (i.e., all levels of the response scale). Hence, if Rasch model results show a separation index of 2.0 for the person measure, we would interpret this as an indication of a screen. In contrast to screens are “measures.” Measures provide the users with the ability to order respondents according to their responses. Generally, the separation index in a Rasch analysis for a measure is 3.0.

Measurement Practice Versus Measurement Development

The purpose of the instrument, as noted, drives the measurement development process. An instrument can be developed through either measurement development or measurement practice [2]. The traditional measurement development process allows the researcher to develop and refine the instrument using a theory-driven process and provide vigilant attention to the process of measurement. In measurement practice, the instrument is used quickly after its development, and these initial data may be the only data regularly collected, so finding evidence of psychometric problems can be quite difficult [2]. Both the VA-13 and the FAST fall into the category of measurement practice. The identified purpose of the VA-13 was to measure the effectiveness of the inpatient blind rehabilitation-training program and outpatient services that

are provided by Blind Rehabilitation Outpatient Specialists (BROSs) to veterans who are not able to attend a BRC. The purpose defined for the FAST was to provide clinicians with a standardized clinical assessment to determine a veteran’s level of visual functioning at admission and discharge (i.e., a measure of efficacy). The FAST was also to be used to measure SWBRC program outcomes [20]. While sufficient attention was given to measurement processes, limitations existed in the development process imposed by the urgent need for national outcomes data and the restrictions of clinical practice. As a result of these limitations and the need to keep respondent burden minimal (i.e., have the fewest number of items possible), both the VA-13 and FAST function as screens.

On the other hand, the VA LV VFQ went through the measurement development process and was developed for measuring the outcomes of low-vision services in clinical research. The instrument developers determined that before a meaningful measure of outcomes could be developed, a clear understanding of the construct that the low-vision intervention attempts to modify was needed. The conceptual framework of the systems model developed by Robert Massof guided development of the VA LV VFQ [24]. This process enabled the instrument developers to address item selection, scaling, and the revisions that were necessary to make the instrument perform as an interval measure rather than a screen.

Content Selection and Generation of Items

Once the purpose of the instrument is established, the development process turns to the content. Surveying the relevant content means deriving an understanding of the breadth and depth of pertinent content. The content for all three instruments is based on aspects of the existing VA blind rehabilitation-training curriculum. The curriculum

comprises the four primary disciplines. The four disciplines in VA blind rehabilitation are—Low Vision, Orientation and Mobility, Living Skills (Rehabilitation Teaching), and Manual Skills. The Manual Skills training, unique to the VA, is designed to assess and enhance skills in all aspects of sensory awareness, with an emphasis on adaptive and safety techniques. The comprehensive training curriculum addresses skill development in activities of daily living (ADL) and instrumental activities of daily living (IADL). The ADL most often addressed include basic activities of hygiene and personal care, such as dressing, grooming, eating, and locomotion or mobility. ADL function is often disrupted by visual impairment; therefore, the activities serve as a gauge for monitoring the effect of rehabilitation interventions [25–26]. IADL are typically assumed to be more complex tasks than ADL, although overlap in complexity is not unusual [27]. IADL cover a much broader range of personal behaviors, such as managing medications, managing money, telling time, shopping for personal needs, housekeeping, and travel outside of the home [28]. With the purpose and content identified, the investigators set out to generate the content domain. The generation of items based on the content domain may be either empirical or theoretical.

Blind Rehabilitation Service Functional Outcomes Survey

In the early 1990s, researchers at the Atlanta VA RR&D Center developed a 32-item clinician-administered blind rehabilitation outcomes instrument, the Functional Independence Measure for Blind Adults (FIMBA) [29–30]. Long et al. developed the FIMBA by adapting items from a previously developed instrument in the 1980s, the Blind Rehabilitation Functional Independence Measure (BRFIM). The BRFIM was an initial attempt to create an assessment tool for blind rehabilitation that was modeled after the Functional Independence Measure (FIM) used in rehabilitation medicine [31]. Clinicians at four BRCs were instructed to administer the FIMBA to veterans at preadmission and postdischarge interviews. Veterans were asked specifically to report on the frequency, ease (or difficulty), and satisfaction with the performance of 32 specific tasks associated with blind rehabilitation. In spite of the researchers' intent to make the FIMBA easy to administer, the interviews and data recordation were unwieldy for clinicians. As a result, an insufficient number of data were collected to establish the reliability and validity of the FIMBA [16].

The items (or questions) for the VA-13 were then selected from the item pool generated in the FIMBA project. Four discipline-specific consensus panels were established, and each included a researcher and a national outcome coordinator (i.e., a designated BRC Director) who arranged and moderated bimonthly conference calls with the discipline supervisors from each of the ten BRCs. Each discipline supervisor worked with his or her individual staff members for ~6 months to identify the items from the item pool that they believed were the best indicators of *functional independence* in that discipline. Feedback from the four consensus panels was assembled and reviewed by members of the BRS Outcomes Task Force. The Atlanta VA RR&D project investigators compiled all recommendations, and this process resulted in the 51-item survey, the BRSFOutSur.

Subsequent field-testing and iterative data-reduction techniques including factor analysis and the identification of high interitem correlations led the instrument developers to reduce the number of items to 28. Next, subject matter experts on the consensus panels were asked to evaluate the 28 items and identify the items they deemed the most salient indicators of independence. This feedback was considered along with additional pilot testing results. A 13-item version was ultimately selected, and the BRSFOutSur instrument was subsequently renamed the VA-13. The VA-13 included questions about the frequency of, independence in, and satisfaction with performing specific tasks [15–16].

Functional Assessment of Self-Reliance on Tasks

The content of the FAST items was determined both theoretically and empirically. Theory dictated that the instrument distinguish between tasks (i.e., specific daily activities) and goals (i.e., general activity categories, where tasks may be grouped) [24]. The primary measurement goal was to determine the change in levels of visual functioning. The instrument developers determined that the first step would require experts (clinicians) to identify both the goals and the specific tasks that visually impaired veterans must be able to accomplish to live independently. Additionally, both empirical and theoretical approaches were used to poll experts for their perspectives on the relevant content domain. The empirical approach involved the application of a modified Delphi procedure, where clinicians were asked questions and all responses were aggregated and sent back to these experts for review and rerating [17]. The theoretical approach

involved a comparison of the Delphi procedure results with the extant literature to ensure some common theoretical convergence, while offering unique data that might be more informative than the traditional theories. The clinicians not only provided “expert” perspectives on content but also feedback about the limitations of instrumentation in the blind rehabilitation-training environment. Great care was also taken to minimize the number of items and lessen the respondent burden for instructors. Once items were generated, they underwent formal pilot testing. Both clinician experts and the instrument developers reviewed all items and assisted in item refinement and definitions. The FAST goal template developed in the previous steps served as the basis for which all items would be written. At least one item was written for each goal. In many circumstances, a single item was sufficient to capture the entire goal domain. However, in several instances, multiple items were necessary to adequately capture the breadth of a goal. The number of FAST items went from 15 to 11 items over a 5-year period, which also included nine rounds of item revisions [20].*

Department of Veterans Affairs Low Vision Visual Functioning Questionnaire

The VA LV VFQ was developed by combining information on patients’ perceptions of needs for low-vision devices with information from rehabilitation professionals on the goals addressed by rehabilitation programs and expected training outcomes [21]. Structured interviews were conducted with 149 patients from the Chicago Lighthouse, Edward Hines Jr. VA Hospital BRC, and the Jessie Brown VA Medical Center Vision Impairment Center to Optimize Remaining Sight to identify the importance of low-vision devices for performing daily living tasks [22]. Need for low-vision devices for close, intermediate, and far distance reading tasks were most frequently reported.

In round 1 of a modified Delphi method, informal focus groups, workshops, and personal interviews were conducted with rehabilitation professionals to obtain their clinical impression of needs and goals of patients with low vision and the rehabilitation programs prescribed for these patients [22]. Low-vision clinical practice

guidelines for optometry, ophthalmology, occupational therapy, and blind rehabilitation were considered in choosing activities for the questionnaire. In round 2 of the modified Delphi method, the information gathered from multiple sources was collated to generate the initial draft of the questionnaire. The Hines BRC research staff field-tested items in individual interviews with veterans to identify problems in wording, item selection, or order.

Generation of Item Responses

Construction of the questions is extremely important but no more important than writing the response options for the questions and thinking about how they will be scored. On the surface, this step appears restricted to multiple choice, forced-response questions (e.g., true-false, check boxes.), and Likert-style items. However, all instruments require some form of response options. Some instruments rely on open-ended responses (e.g., semistructured interviews), while other instruments contain forced responses (e.g., multiple-choice options). The characteristics that make good response options for performance measures include (1) an undeniably correct response, (2) clear language that allows respondents to discriminate among the choices, and (3) distracters that appropriately modulate the difficulty of the question. These issues are especially important given that the self-report versions of instruments must be interviewer-administered to the visually impaired [2].*

Blind Rehabilitation Service Functional Outcomes Survey

In addition to generating content-relevant items, the four discipline-specific consensus panels recommended a multidimensional response set that was based on determinants (i.e., patients’ needs and interests) to account for the expected outcomes (i.e., increased frequency of activities, independence, and satisfaction). A VA-13 trained interviewer asks a veteran to rate his or her current ability on the 3-point scale by indicating whether the task could be completed with “1—a great deal of assistance,” “2—a little assistance,” or “3—no assistance (independently).” Attention was given to the feasibility of administering the instrument in terms of its length and potential for respondent burden. The VA-13 scale properties were evaluated in a sample of 190 veterans for both administrations of the scale [20]. An extreme distribution of scores for this sample indicated that most patients responded to the items using primarily one side of the rating scale. The 3-point scale functions as a dichotomous scale with most

*McKnight PE, Babcock-Parziale JL, Head DN. Ten essential steps to developing a blind rehabilitation outcome measure. Unpublished observations, 2006.

of the patient ratings occurring at responses 1 and 3 for both the retrospective pre- and posttests.

Functional Assessment of Self-Reliance on Tasks

The response options for the FAST were determined largely by the nature of the measurement task but also influenced by the respondents (clinicians). In clinical practice, the purpose of the FAST is to assess a patient's level of visual functioning (or functional independence) before and after vision rehabilitation services. Clinicians, therefore, emphasized that identifying the levels of independence in the descriptions of each response choice was important. The first version of the FAST made use of five response values, which contained anchors of extreme, severe, moderate, slight, and no problems. Essentially, these response choices were ratings of the amount of assistance required for each goal [17]. The response options were initially influenced by the ratings used for the FIM [32], but over time, the options were changed to conform to the goals of the measurement task. For the next two revisions, the 5-point response options were changed because the clinical staff was dissatisfied with the response anchors as they were initially worded. A 10-point scale was eventually developed for version 4 and used thereafter. The 10-point scale resolved many of the problems inherent in the 5- and 6-point scales and was the set of response choices that the clinicians favored. The wording changed repeatedly for the 10-point anchors in each version to increase the clarity of the rating response descriptors. The process used to refine the response options helped both clinicians and researchers better interpret the information gained by the items [2].* The item response categories have been tested in a previous analysis of over 500 FAST cases.† All categories of the 10-point scale are endorsed; however, the lower categories (1–4) are endorsed most often at pretest and the higher categories (7–10) are endorsed more often at posttest. While the item cluster is indicative of a screen, the rating scale provides enough variability to differentiate the items sufficiently and allow reasonable judgments about successful attainment of the

functional goals associated with low-vision and blind rehabilitation [20].

Department of Veterans Affairs Low Vision Visual Functioning Questionnaire

Identical response scales were used for all items to simplify the task of completing the survey. Activities performed at far, intermediate, and near distances both at home and in the outside community were included. For each item, patients were asked, "Is it difficult to...?" The 5-point response scale included the following options: not difficult, slightly difficult, moderately difficult, extremely difficult, or impossible. The difficult/don't do for nonvisual reasons response option was scored as missing data. Three additional questions were included following pilot testing. "Is it (difficult) because of your vision?" Response choices were yes and no. "Do you want training?" Response choices were yes or no. Patients who were able to perform the activities were asked, "How do you usually ...?" (perform the activity). Response choices were (1) own eyes or eyeglasses, (2) vision devices/techniques (e.g., magnifier), (3) other senses/nonvisual devices (e.g. cane), (4) someone helps me, and (5) not applicable. To determine the range and precision of the instrument for measuring visual ability of low-vision patients with moderate-to-severe vision loss across diverse clinical settings, researchers obtained pre-rehabilitation telephone interviews from 367 subjects at five centers with different low-vision services. Subjects at all five centers used the rating categories the same way. The ratings of slight and moderate difficulty were used interchangeably, which suggested that the response scale could be reduced to four choices. Researchers administered the questionnaire to 50 additional subjects to confirm that the scale was analyzed in the same way when four rather than five difficulty ratings were listed (not difficult, slightly/moderately difficult, extremely difficult, impossible) [22].

Pretesting, Revising, and Field-Testing the Instruments

Once the item content and response option formats are reasonably established, pretesting the questions to ensure that respondents understand the questions, offer valid responses, and complete all appropriate questions is wise [2].* Pretesting often refers to a "trial run" administration of a measure followed by a debriefing period. This is an acceptable view of pretesting; however, pretesting constrained to this method provides only limited

*McKnight PE, Babcock-Parziale JL, Head DN. Ten essential steps to developing a blind rehabilitation outcome measure. Unpublished observations, 2006.

†Babcock-Parziale JL, Head DN, McKnight PE, Massof RW. Validation of the Functional Assessment of Self-reliance on Tasks. Invited paper presented at the American Academy of Optometry; 2002 Dec; San Diego, CA.

information. Pretesting ought to involve not only a representative sampling of appropriate respondents but also content experts in the process. The respondents provide information about readability, and the experts provide feedback about more subtle aspects of the instrument. Specifically, content experts may offer insights into the clustering of questions, response choice options, and the general design of the instrument. Respondents may not necessarily be knowledgeable enough to provide this feedback. The data gleaned from those who participate in pretesting may be informative for altering the instrument or serve as a confirmation that the instrument is sufficiently mature and ready for use. Obviously, if pretesting data provide critical feedback about the design of any facet of the measure, steps must be taken to rectify the situation and employing an additional pretesting occasion may be useful.

Although each of the three instruments was pretested, the amount of pretesting varied across the instruments. The VA-13, for instance, was pretested to clarify the respondents' understanding of the items for the interviewer-administered format. However, once the instrument began to be administered nationally to all veterans who completed an inpatient stay at one of the ten BRCs, the instrument items and scaling were not revised.

As noted previously, the FAST went through nine iterations between 1997 and 2002. The first five iterations were essentially pretests, followed by four field-tests of the instrument [20].*

The VA LV VFQ went through two pretesting rounds prior to field-testing. In the third round, the low-vision staff, investigators, and research coordinators made a final selection of the items included in the initial field-test version of the instrument [22]. Items used on the 48-item field-test version of the VA LV VFQ represent tasks that are difficult to perform because of loss of visual acuity, visual field, and/or contrast sensitivity deficits. Although similar items are found on many visual functioning questionnaires, the item descriptions were condensed to be user-friendly for seniors.

Administration Time and Approach

Blind Rehabilitation Service Functional Outcomes Survey

The instrument is administered by a trained research staff member to veterans as a posttreatment telephone survey 4 to 6 weeks following discharge from any one of the 10 BRCs. At the time of the posttreatment interview, the veteran is also asked to rate his or her current ability on the 3-point scale on each of the 13 tasks that address independence. Following the report of current ability (posttest), the veteran is then asked to rate his or her ability to perform the same task prior to enrolling in blind rehabilitation (a retrospective pretest). Administering the 13 items for both the posttest and retrospective pretest requires approximately 15 to 20 minutes.

Functional Assessment of Self-Reliance on Tasks

The FAST developers chose to administer all items via a paper-and-pencil medium and have the clinicians be the instrument users. Instructions were therefore written in a language that clinicians easily understand. Moreover, the descriptors for each response option provided exemplars so that the clinicians might anchor their selections to characteristics that they observe routinely in their daily lessons. Clinicians use the FAST to rate veterans at admission and just prior to discharge. Some researchers have argued that clinician ratings are more biased because of an evaluation phobia (i.e., they score veterans low at pretest and high at posttest to maximize the program effects or outcomes). Teams, however, rate the FAST, and any one clinician feeling threatened by the evaluation process is unlikely. Furthermore, no empirical evidence exists to document that clinician bias or error is any greater than the error that is associated with self-reports [19–20]. Completing the FAST takes a clinical team approximately 15 minutes per patient.

Department of Department of Veterans Affairs Low Vision Visual Functioning Questionnaire

The VA LV VFQ is interviewer-administered to patients prior to enrollment in either outpatient low-vision services or an inpatient blind rehabilitation program. Follow-up interviews are usually conducted 6 to 8 weeks after discharge. Administration time for all questions on this instrument ranges from 25 to 35 minutes [22].

*McKnight PE, Babcock-Parziale JL, Head DN. Ten essential steps to developing a blind rehabilitation outcome measure. Unpublished observations, 2006.

Overview of Psychometric Analyses Conducted

Most instrument development efforts are limited by timing or funding, which may restrict the opportunity to conduct meaningful psychometric analyses. A new instrument is traditionally analyzed using classical test theory (CTT). See Massof for an advanced discussion on CTT and the measurement of vision disability [1]. Instrument developers will most often analyze the raw scores, create a simple scoring algorithm (i.e., sum the raw item scores), and report very basic statistics. The prevailing paradigm in clinical vision research, for instance, was to synthesize instrument scores from Likert scaling of patient ratings, to accept instrument scores as measurements, and to validate the claims of measurement with inferential evidence. The validation evidence for questionnaires often consists of correlations of instrument scores with external variables (e.g., visual acuity), demonstrations of “acceptable” values of Cronbach alpha to demonstrate reliability, and demonstrations that instrument scores could be used to discriminate groups of respondents who were categorized by other criteria [1,12].

The new rules of measurement, however, specify that to qualify as a measure, an instrument must produce a ratio or interval scale [1–2,12–13]. VA vision rehabilitation investigators followed this recommendation and used Rasch probabilistic conjoint measurement models to estimate measurements from respondents’ ratings of items. The development of the FAST and the VA LV VFQ, for example, relied heavily on Rasch analysis and used the intensive item analysis approach [14,20–22]. Developers of the VA-13 relied initially on CTT analyses conducted on the raw scores [15–16], and although a Rasch analysis was later conducted, these data were not published. In 2003, a Rasch analysis of the VA-13 was conducted in conjunction with the FAST on a sample of 190 veterans [20]. Investigators initially planned to calibrate the two instruments, but the Rasch analyses revealed irregularities in both instruments that precluded calibration efforts. Analyzing both instruments with the same sample of veterans, however, highlighted the important similarities and differences between the instruments (**Table 2**). Recognizing the role these differences play in the measurement process was instrumental in helping the investigators define future development efforts. The pertinent psychometric findings for the three instruments are summarized below.

Psychometric Findings: Blind Rehabilitation Service Functional Outcomes Survey

Rasch model analyses of the VA-13 retrospective pretest and posttest were performed on a sample of 190 veterans who participated in vision rehabilitation at the SWBRC. The analysis identified three critical issues that require further attention and development effort. The retrospective pretest revealed a strong ceiling effect for this sample, similar to that seen for previous samples [20]. A ceiling effect occurs when many subjects in a study have scores that are at or near the possible upper limit or “ceiling.” The ceiling effect makes analysis difficult because it reduces the amount of variation and restricts the range observed in the scores. Second, the VA-13 demonstrated poor sensitivity to change—a problem that may be due to the aforementioned ceiling effect. Third, in addition to the ceiling effect that occurs for the retrospective pretest and posttest scores, the restricted use of the rating scale may account for the lack of variability and the minimal difference in observed change in the VA-13. As a result, the 3-point scale does not appear to capture the full variability in patients’ visual functioning that it was intended to capture. Based on these findings, the VA-13 currently does not meet the measurement criteria specified in the “Methods” section. The restricted scale and limited variability across items prohibit the discrimination within patients and between patients over time.

Psychometric Findings: Functional Assessment of Self-Reliance on Tasks

Rasch results for the FAST, based on the same sample of 190 veterans, indicate that the person estimates change reliably from pretest to posttest; however, the instrument requires additional revisions [20]. First, the FAST has two items (items 2 and 3) that appear to become more difficult following rehabilitation. These two items must be revised to ensure that their difficulty levels remain consistent over time. The FAST item that measures “Reading” becomes easier at posttest. This item must be revised to account for the change in the context of reporting (caused by the introduction of reading aids) and to improve the Rasch model fit of this item. The findings indicate that the FAST fails to meet the first measurement criteria because the current items are unable to detect change at the extreme lower and upper ends of disability. Second, the FAST results reveal that the measure captures the expected shift from unable to able to complete the tasks over the course of the blind rehabilitation

Table 2.

Similarities and differences between instruments.

Instrument	Item Level	Respondent	Number of Items	Rating Scale	Time of Administration	
					Pretest	Posttest
FAST	Goal	Clinician-Rated	10	10-point	Admission	Discharge
VA-13	Task	Self-Report	13	3-point	Retrospective Pretest	Postdischarge
VA LV VFQ	Task	Self-Report	53	4-point	Admission	Postdischarge

FAST = Functional Assessment of Self-Reliance on Tasks, VA-13 = Blind Rehabilitation Service Functional Outcomes Survey, VA LV VFQ = Department of Veterans Affairs Low Vision Visual Functioning Questionnaire.

program. However, a ceiling effect exists in the posttest administration of the FAST, just as it did for the posttest administration of VA-13. The FAST accounts for little variability at posttest (discharge), when veterans tend to reach a similar level of functioning. Despite the ceiling effect, the Rasch model results indicate the FAST offers a separation index greater than 2.0 and, therefore, reliably differentiates between successful and unsuccessful rehabilitation outcomes. (This finding satisfies the second measurement criterion.)

Psychometric Findings: Department of Veterans Affairs Low Vision Visual Functioning Questionnaire

The most recent psychometric properties published for the VA LV VFQ are based on a sample of 367 patients with visual acuity ranging from near normal to total blindness at three VA sites and two private sector sites [21]. These results are limited to the difficulty rating of items obtained before rehabilitation. Findings indicate that 7 of the 48 items did not fit Rasch model requirements and may need to be rewritten [22]. The instrument developers acknowledge that some items could be omitted to shorten the instrument and decrease administration time and respondent burden. However, we concur that reducing the number of items before the sensitivity of the items to change can be evaluated is premature. Preliminary results of the instrument's ability to detect change were reported recently at a national meeting [33]. The measure was administered 1 week before and 3 months following discharge to 220 subjects (134 subjects from the outpatient low-vision clinics and 86 subjects from the inpatient Hines BRC program). The VA LV VFQ was able to detect change both within and between persons. The effect size (the magnitude of the effect) for the two patient groups was 1.45 for the inpatient group and 0.18 for the outpatient group. The smaller effect realized for patients in the outpatient programs is somewhat expected

because these programs accept all patients with visual impairments ranging from mild to profound, while the inpatient program accepts only those patients who are legally blind. Hence, low-vision outpatients that enter with less severe visual impairment will be less likely to show as much change in functional ability compared with those in the inpatient program. The smaller difference accounted for between programs in postrehabilitation visual ability also suggests the possibility of a rehabilitation ceiling effect similar to that found for the VA-13 and the FAST (e.g., patients cannot be rehabilitated to levels of visual ability that exceed the premorbid state).

These findings suggest that the VA LV VFQ meets both of the measurement criteria. However, the current version of the instrument is quite lengthy and increases respondent burden. In the near future, the instrument developers will need to select a smaller item pool that can assess program outcomes across the continuum of VA low-vision and blind rehabilitation services. The condensed version will require a new evaluation to determine if the instrument meets the suggested criteria.

DISCUSSION

This article provides a retrospective account of the development and evaluation of the three measures. To date, both the quality and type of data available for each of the instruments varies, which makes the psychometric findings quite different and the direct comparison of these findings difficult at best. As noted, the VA-13 was developed first and has been used to measure the VA BRS inpatient and BROS outpatient training program outcomes. The two measures that followed, the FAST and the VA LV VFQ, have taken a different developmental approach than the VA-13 at multiple levels. As the three measures progressed through various developmental

stages, the fact that an objective evaluation of the three instruments was not possible became apparent because no universal assessment criteria had been selected for the field. Under ideal conditions, the instruments ought to be evaluated with the same methods and held to the same standards. A comparison of the FAST and the VA LV VFQ became possible once both developers used a similar analytic approach, Rasch analysis. The analytic approach provided a common method for comparing results; therefore, the same approach was later used to examine the VA-13. The findings from this analytic approach provide evidence about how well each of the instruments met the specified psychometric criteria and provided specific information that was used to development efforts.

While this review provides a historical overview of the instrument development efforts, the field will not advance in its measurement of low-vision and blind rehabilitation outcomes without designated measurement standards that researchers consistently apply to develop and evaluate all instruments. To this end, we want to reiterate the two criteria we initially specified and suggest a third for future consideration.

First, a low-vision and blind rehabilitation outcomes instrument must reliably discriminate between the levels of visual functioning for patients in this population (from legal blindness to no light perception). The goal is to capture the broadest array of functional domains possible, without decreasing the sensitivity to change within the domains. Instruments that meet these criteria will be able to measure the change in patients' visual functioning across the continuum of services provided by the VA. Second, an instrument must be sensitive to or able to detect change across time for all levels of visual functioning, both within and between patients. Meeting these criteria is essential—but not sufficient—for good measurement.

A third and important criterion is that instruments developed in this field be sensitive to clinically relevant effects. The central issue of rehabilitation is the ability to change performance; therefore, evaluating whether these changes are clinically significant (i.e., whether they make any real functional difference to patients in their everyday lives) is important [34]. If a measure meets the first two criteria and reliably discriminates across the broadest array of functional domains, ideally it should also provide clinicians with relevant information for treatment planning and program development. As part of the evolution of outcomes measures, VA instrument developers

have already begun to integrate these measurement criteria into their instrument developmental goals.

Development Goals: Blind Rehabilitation Service Functional Outcomes Survey

Instrument developers at the Atlanta RR&D Center provided several development goals generated by both empirical and collegial feedback. As noted, the Rasch analysis revealed that the original 3-point response scale functioned as a dichotomous scale. The pronounced ceiling effect in both retrospective pretest and posttest scores minimized both item variability and the total variance accounted for by the instrument. In response, the VA-13 development team began field-testing a new version of the instrument that includes a 4-point response scale for the independence domain. These data will be analyzed to determine if the new scale functions as an interval measure and is able to detect change across the two time points.

An additional concern expressed about the VA-13 was the exclusive use of the retrospective pretest in lieu of a traditional pretest. The retrospective pretest is warranted only if a response shift is demonstrated in the self-report ratings [35]. An important development goal, therefore, involves the collection of true baseline assessments of veterans' visual functioning. Using the revised VA-13, researchers are currently conducting interviews with veterans before admission to and following discharge from a BRC. The retrospective pretest is also administered as a part of the follow-up interview. Once sufficient samples of both pretest and retrospective pretest data are collected, the data will be analyzed to determine if a response shift is noted in the veterans' ratings. After findings are reviewed, investigators will empirically establish the best pretest method for administering the VA-13.

During this intermediate stage of instrument development, the investigators are administering the VA LV VFQ and the VA-13 to about 200 subjects. A Rasch analysis will then be performed on data for both instruments using the same subjects. An intensive item analysis will be conducted, similar to the one conducted with the FAST and the VA-13 [20]. This analytic approach will allow investigators to better understand item measurement at the task level for both self-report instruments. Investigators believe the current revisions to the VA-13 and the application of subsequent empirical findings will further enhance the instrument's performance and ensure that it meets the specified measurement criteria.

Development Goals: Functional Assessment of Self-Reliance on Tasks

Developers of the FAST are addressing the noted empirical concerns. Currently, the CR and a newly developed SR version of the FAST are used to monitor SWBRC program outcomes. In addition to clinician ratings at admission and discharge, the SR version has been administered to veterans prior to admission, at discharge, and at 3-months postdischarge [36]. Researchers have rewritten several of the items on both versions in an attempt to maximize item difficulties. The new versions of the instruments will be field-tested to determine if the revised items fit the Rasch model and alter the FAST from a screen to an interval measure.

Investigators at the SWBRC contend that establishing the construct validity for each of the three instruments is a systematic and necessary process. They have been funded by the VA Health Services Research and Development Service to conduct a 4-year study, "Advancing outcomes measurement and microcosting in blind rehabilitation." The study employs a fully crossed design to determine whether the three measures account for the same functional outcomes given the differences across the instruments identified in **Table 2**. First, the three measures differ in terms of emphasis—the FAST emphasizes functional ability, while the VA-13 and VA LV VFQ emphasize functional independence. Second, the measures differ by respondent. The unique variance measured by both perspectives (clinician and self) will be investigated to determine just how important the two perspectives may be in terms of clinical relevance and documenting outcomes. Determining the relative importance of respondent and emphasis would help clarify the construct and provide better interpretation of vision rehabilitation program outcomes. Findings from this study should advance the theory of outcomes measurement for the field by addressing the biases associated with respondents and determining the construct validity of the instruments.

Development Goals: Department of Veterans Affairs Low Vision Visual Functioning Questionnaire

The VA LV VFQ is currently used as an outcomes measure in two low-vision clinical trials funded by the VA RR&D Service, the "VA low-vision intervention trial" and "Measuring reading rehabilitation outcomes," as well as a study, "Development of methods and tools to evaluate eccentric viewing training." Currently, researchers at the Atlanta RR&D Center are using the VA LV

VFQ 53-item version in the study, "Caregiver burden and the rehabilitation of aging visually impaired." Data on 1-year follow-up of veterans in outcomes studies at the Hines BRC are currently being analyzed and compared with data from 3-month follow-up to assess stability of the instrument. Investigators are calibrating new items for the VA LV VFQ and refining the calibration of existing items through additional data collection and analyses. The long-term goal for further development of the VA LV VFQ is to calibrate a data bank of items that can be used for the response question on "Difficulty performing activities." Items from the data bank can then be chosen depending on the research question asked and the visual ability of subjects. Use of computer-adaptive testing, in which tests are tailored to the visual ability level of the patient, will facilitate this process.

CONCLUSIONS

Developing measures to assess low-vision and blind rehabilitation outcomes is not an end in itself but rather the beginning of a new evidence base to inform practice. Each of the instruments reviewed in this article was developed to account for patients' outcomes. Ultimately, the primary purpose for collecting outcomes data should be to establish an evidence base to inform practice and maximize the efficiency and effectiveness of the services provided to visually impaired veterans. Although none of the measures reviewed currently meets all aspects of the proposed measurement criteria, we hope that future development efforts both within and outside the VA will incorporate these criteria. The development of instruments that meet the specified measurement standards and demonstrate clinical relevance is an essential step toward establishing this evidence-based system for the VA BRS and the field of low-vision and blind rehabilitation.

ACKNOWLEDGMENTS

We wish to acknowledge and thank the veterans who participated in these studies; without their participation, scientific and clinical knowledge would not advance. We want to thank Patrick McKnight, PhD, of George Mason University, for his insightful and informative review of this article. We appreciate his collaboration and invaluable expertise in measurement. We also want to thank

Daniel N. Head, EdD, of the SWBRC, for his support and input in the preparation of this manuscript, and Joan A. Stelmack, PhD, OD, of the Edward Hines Jr. VA Hospital BRC, for providing current research findings for the VA LV VFQ.

This material was based on work supported by the VA, Veterans Health Administration, RR&D Service (De l'Aune and Williams, grants C1776R, C1894-R, and C2269-R; Babcock-Parziale, grant C-2710I; and Stelmack, grants 02086R and C2707I).

The authors have declared that no competing interests exist.

REFERENCES

1. Massof RW. The measurement of vision disability. *Optom Vis Sci.* 2002;79(8):516–52. [PMID: 12199545]
2. Wilson M. *Constructing measures: An item response modeling approach.* Mahwah (NJ): Lawrence Erlbaum; 2004.
3. Stelmack JA, Stelmack TR, Massof RW. Measuring low-vision rehabilitation outcomes with the NEI VFQ-25. *Invest Ophthalmol Vis Sci.* 2002;43(9):2859–68. [PMID: 12202503]
4. Goodrich GL. *Growth in a shrinking population: Visual impairment in the veteran population, 1995–2010.* Palo Alto (CA): VA Palo Alto Health Care System, Western Blind Rehabilitation Center; 1995.
5. Kizer KW. *Prescription for change: The guiding principles and strategic objectives underlying the transformation of the Veterans Healthcare System.* Washington (DC): Department of Veterans Affairs; 1996.
6. Margolis MK, Coyne K, Kennedy-Martin T, Baker T, Schein O, Revicki DA. Vision-specific instruments for the assessment of health-related quality of life and visual functioning: A literature review. *Pharmacoeconomics.* 2002; 20(12):791–812. [PMID: 12236802]
7. Raasch TW, Leat SJ, Kleinstein RN, Bullimore MA, Cutter GR. Evaluating the value of low-vision services. *J Am Optom Assoc.* 1997;68(5):287–93. [PMID: 9170794]
8. Mangione CM, Berry S, Spritzer K, Janz NK, Klein R, Owsley C, Lee PP. Identifying the content area for the 51-item National Eye Institute Visual Function Questionnaire: Results from focus groups with visually impaired persons. *Arch Ophthalmol.* 1998;166(2):227–33. [PMID: 9488276]
9. Mangione CM, Lee PP, Pitts J, Gutierrez P, Berry S, Hays RD. Psychometric properties of the National Eye Institute Visual Function Questionnaire (NEI-VFQ). NEI-VFQ Field Test Investigators. *Arch Ophthalmol.* 1998;166(11): 1496–1504. [PMID: 9823352]
10. Stelmack JA, Babcock-Parziale JL, Head DN, Wolfe GS, Fakhoury NE, Wu SM, Massof RW. Timing and directions for administration of questionnaires affect outcomes measurement. *J Rehabil Res Dev.* 2006;43(6):809–16.
11. Stelmack JA, Szlyk JP, Joslin C, Swetland B, Myers L. Pilot study: Use of the NEI VFQ-25 to measure outcomes of low vision rehabilitation services in the Department of Veterans Affairs. In: Stuen C, Arditi A, Horowitz A, Lang MA, Rosenthal B, Seidman K, editors. *Vision rehabilitation (Assessment, intervention and outcomes).* Exton (PA): Swets & Zeitlinger; 2000. p. 774–76.
12. Massof RW, Fletcher DC. Evaluation of the NEI visual functioning questionnaire as an interval measure of visual ability in low vision. *Vision Res.* 2001;41(3):397–413. [PMID: 11164454]
13. Wright BD, Linacre JM. Observations are always ordinal; measurements, however, must be interval. *Arch Phys Med Rehabil.* 1989;70(12):857–60. [PMID: 2818162]
14. Stelmack JA, Szlyk JP, Stelmack TR, Babcock-Parziale JL, Demers-Turco P, Williams TR, Massof RW. Use of Rasch person-item map in exploratory data analysis: A clinical perspective. *J Rehabil Res Dev.* 2004;41(2):233–41. [PMID: 15558377]
15. De l'Aune WR, Williams MD, Welsh RL. Outcome assessment of the rehabilitation of the visually impaired. *J Rehabil Res Dev.* 1999;36(4):273–93. [PMID: 10678451]
16. De l'Aune WD, Watson GR, Schuckers P, Ventimiglia G. Clinical application of a self-report, functional independence outcomes measure in the DVA's Blind Rehabilitation Service. *J Vis Impair Blind.* 2004;98(4):197–211.
17. Head DN, Babcock JL, Goodrich GL, Boyless JA. A geriatric assessment of functional status in vision rehabilitation. *J Vis Impair Blind.* 2000;94(6):357–71.
18. Babcock JL, Goodrich GL, Head DN, Boyless JA. Developing geriatric training outcome assessments in vision rehabilitation. *J Vis Impair Blind.* 2000;94(5):307–21.
19. Babcock JL, Head DN, McKnight PE, Massof RW. Establishing benchmarks for blind rehabilitation by calibrating clinical and self-report measures. In: *Proceedings of the VA Health Services Research and Development Service 20th Annual Meeting; 2002 Feb 14–15; Washington (DC): Department of Veterans Affairs; 2002.* p. 129–30.
20. Babcock-Parziale JL, McKnight PE, Head DN. Evaluating psychometric properties of a clinical and a self-report blind rehabilitation outcome measure. *J Rehabil Res Dev.* 2005; 42(4):487–98. [PMID: 16320144]
21. Szlyk JP, Stelmack JA, Massof RW, Stelmack TR, Demers-Turco P, Williams RT, Wright BD. Performance of the Veterans Affairs Low Vision Visual Functioning Questionnaire. *J Vis Impair Blind.* 2004;98(5):261–75.
22. Stelmack JA, Szlyk JP, Stelmack TR, Demers-Turco P, Williams RT, Moran D, Massof RW. Psychometric properties

- of the Veterans Affairs Low Vision Visual Functioning Questionnaire. *Invest Ophthalmol Vis Sci.* 2004;45(11):3919–28. [PMID: 15505037]
23. Lord FM. Applications of item response to theory to practical testing problems. Hillsdale (NJ): Lawrence Erlbaum; 1980.
24. Massof RW. A systems model for low vision rehabilitation. I. Basic concepts. *Optom Vis Sci.* 1995;72(10):725–73. [PMID: 8570162]
25. Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW. Studies of illness in the aged. The index of ADL: A standardized measure of biological and psychosocial function. *JAMA.* 1963;185:914–19. [PMID: 14044222]
26. Granger CV, Mann WC, Ottenbacher KJ, Tomita MR, Fiedler RC. Functional measures of geriatric subjects in the community. *Top Geriatr Rehabil.* 1994;10(1):7–21.
27. Spector WD, Fleishman JA. Combining activities of daily living with instrumental activities of daily living to measure functional disability. *J Gerontol B Psychol Sci Soc Sci.* 1998;53(1):S46–57. [PMID: 9469176]
28. Branch LG, Hoenig H. Measures of physical functioning. *Generation.* 1997;21(1):37–40.
29. Crews JE, Long RG. Conceptual and methodological issues in rehabilitation outcomes for adults who are visually impaired. *J Vis Impair Blind.* 1997;91(2):117–130.
30. Long RG, Crews JE, Mancil R. Creating measures of rehabilitation outcomes for people who are visually impaired: The FIMBA project. *J Vis Impair Blind.* 2000;94(5):292–306.
31. Keith RA, Granger CV, Hamilton BB, Sherwin FS. The functional independence measure: A new tool for rehabilitation. *Adv Clin Rehabil.* 1987;1:6–18. [PMID: 3503663]
32. Granger CV, Hamilton BB, Linacre JM, Heinemann AW, Wright BD. Performance profiles of the functional independence measure. *Am J Phys Med Rehabil.* 1993;72(2):84–89. [PMID: 8476548]
33. Stelmack JA. Effectiveness of low vision rehabilitation programs [abstract]. In: Proceedings of the American Public Health Association's 132nd Annual Meeting and Exposition; 2004 Nov 6–10; Washington, DC. Washington (DC): APHA; 2004.
34. Kazdin AE. Research design in clinical psychology. Needham Heights (MA): Allyn & Bacon; 1997. p. 381.
35. Schwartz CE, Sprangers MA. Adaptation to changing health: Response shift in quality-of-life research. Washington (DC): American Psychological Association; 2000.
36. Babcock-Parziale JL, Cunningham V. Understanding change in self-perceived functional status in veterans following low vision rehabilitation [abstract]. In: Proceedings of the American Academy of Optometry; 2003 Dec 4–7; Dallas, TX. Rockville (MD): American Academy of Optometry; 2003.

Submitted for publication July 15, 2005. Accepted in revised form January 27, 2006.